APRIL/MAY 2025

23PCS21 — DATA MINING AND WAREHOUSING

Time : Three hours Maximum : 75 marks

SECTION A — (10 × 2 = 20 marks)

Answer ALL the questions.

Sharanga Industries Arts & Science College T.V.Malai

1. What are Similarity measures?
2. Define the term — Genetic algorithm.
3. What is Regression?
4. What is Supervised learning?
5. What are Incremental rules in Association rule mining?
6. What is Apriori algorithm?
7. Define the term — OLAP Systems.
8. What is OLTP Systems?

9. How to build a Data warehouse?

10. Mention any two Organizational issues in Data warehouse design.

SECTION B — (5 × 5 = 25 marks)

Answer ALL the questions.

11. (a) Describe about Data mining issues and Data mining metrics.

Or

(b) Discuss on the Data mining tasks — Regression, Time series analysis and Sequence discovery.

12. (a) Explain the Distance-based algorithms in Classification.

Or

(b) Describe about Bayesian Classification algorithm.

13. (a) Discuss on Hierarchical algorithms in Clustering.

Or

(b) Explain the Parallel and Distributed algorithms in Association rule mining.

14. (a) Describe the basics of Data Marts.

Or

(b) Discuss on Star Schema for Multidimensional View in Data modeling.

15. (a) Explain the Design considerations, Data content and Meta data in Data warehouse.

Or

(b) Describe about National Data warehouses.

Shanmuga Industries Arts & Science College T.V.Malai

SECTION C — (3 × 10 = 30 marks)

Answer any THREE questions.

16. Explain the working of Neural networks in detail.

17. Describe the Decision tree-based algorithms in Classification.

18. Discuss on Advanced Association rule techniques in detail.

19. Write a detailed note on State of the Market in OLAP Tools.

20. Explain the various Technological considerations in Data warehouse development

———————